URN: 11D/674

The effect of uncertainty on the 2008 fuel poverty statistics

1. Executive Summary

The national estimates for the number of fuel poor households in England are calculated using data from a number of sources. These include the English Housing Survey (formerly the English House Condition Survey) and DECC's Domestic Fuels Inquiry. Each of these inputs is subject to a degree of uncertainty and estimation. For example, the fuel prices used are mean fuel prices for a particular region (not the actual prices paid by each household) and the values of some of the aspects of income are provided as banded estimates, rather than more precise point values. By estimating the uncertainty in each of the inputs into the fuel poverty calculation we are able to approximate a corresponding level of uncertainty associated with the final national fuel poverty figures. This has been done using a multiple, repeated sampling technique known as a 'Monte Carlo' analysis, and the results are reported below.

This report looks solely at income and price uncertainty. Analysis covering all three main elements of uncertainty, income (from estimates of the amount of benefit received), prices (from assumption of the mean tariff), fuel consumption (dwelling factors due to inter-surveyor variability) was undertaken for 2006 and published in the Annual Report on Fuel Poverty Statistics 2009 (see: http://www.decc.gov.uk/Media/viewfile.ashx?FilePath=Statistics\fuelpoverty\1_20091020153241_e_@@_annualreportfuelpovertystats2009.pdf&filetype=4&minwidth=true).

The published estimate for the total number of fuel poor households in England in 2008 is ~3.335 million. The addition of uncertainty in income and fuel prices, acting together, acts to increase the number of fuel poor households very slightly to a 'most likely value' of ~3.343 million households. The lower and upper 95% confidence bounds for the total number of fuel poor households ~3.299 and ~3.388 million households respectively (a range of ~ 88,000 households).

Estimates of the individual effects of uncertainty in income and fuel prices acting on their own have also been produced. Income uncertainty alone results in a most likely value of ~3.338 million fuel poor households with a 95% confidence interval between 3.308 and 3.369 million households (a range of ~ 61,000 households). Fuel price uncertainty acting alone results in a most likely value of ~3.336 million households with a 95% confidence interval between 3.295 and 3.377 million households (a range of ~ 81,000 households).

Contents

2. Preface

Fuel poverty statistics are compiled from modelled survey data. The main sources of data and information come from:

- The English Housing Survey (EHS) – provides information on households and dwellings in England, from which fuel requirement and household income can be modelled.
- Fuel prices from DECCs Domestic Fuel Inquiry – this is a survey of energy suppliers and provides a near census of domestic gas and electricity prices. These are then combined with modelled fuel requirements to produce a modelled bill for each household in the survey.

Other sources of prices are used for non metered fuels – for more information, see section 4.2.

While the methodology underlying the fuel poverty modelling is constantly reviewed to ensure it is robust and appropriate, it is necessary to make a number of assumptions and to use averages in the modelling. DECC periodically considers the impact of these assumptions and the uncertainty that is inherent in making them. Prior to this report, the analysis was most recently undertaken for 2006 data, and is summarised in the annual report on fuel poverty statistics 2009, chapter 5: http://www.decc.gov.uk/publications/basket.aspx?FilePath=Statistics%5cfuelpoverty%5c1_20091020153241_e_%40%40_annualreportfuelpovertystats2009.pdf&filetype=4.

This report, based on 2008 data, was prepared by BRE for DECC and has been extended to look at input uncertainties separately, and to consider the uncertainty amongst various household groups individually.

3. Introduction

A fuel poor household is defined as one which needs to spend over 10% of its annual income on household fuel use in order to maintain an adequate standard of warmth. The 2008 estimate of the total number of fuel poor households in England is around 3.335 million households (approximately 15.6% of all households). Of these, around 2.650 million households (approximately 12.4%) are considered to be vulnerable.

This estimate is produced using a combined dataset collected as part of the English House Condition Survey in 2007 and English Housing Survey in 2008 (EHCS / EHS) together with DECC information on fuel prices. To a greater or lesser degree, each of the inputs into the calculation of fuel poverty are themselves best estimates or approximations and the modelling process itself requires numerous assumptions to be made throughout. As a result, the process incurs an amount of uncertainty. For example, uncertainty in fuel prices arises from the use of a mean annual value, and uncertainty in income can result from data collected as banded amounts, rather than more precise point estimates. These input uncertainties inevitably lend an

uncertainty to any statistic produced using the data. This study attempts to provide an estimate of the magnitude of this uncertainty, and how it affects the estimate of fuel poverty in England in 2008.

4. Methodology Summary

A household is considered to be in fuel poverty if, in order to maintain an adequate standard of warmth, it needs to spend more than 10% of its income on all household fuel use[1]. This criterion is expressed quantitatively for any particular household by means of the fuel poverty ratio (FPR):

$$\text{Fuel poverty ratio} = \frac{\sum_{\text{All fuels}} (\text{Unit Fuel Price} \times \text{Fuel Consumption}) + \sum_{\text{All fuels}} (\text{Standing charge})}{\text{Income}} \quad \text{(Eqn 1)}$$

Any household with an FPR greater than 0.1 is considered to be in fuel poverty (i.e. the household needs to spend more than 10% of its income on household fuel use to achieve an adequate standard of warmth).

The number and proportions of fuel poor households are obtained by calculating the FPR for each household in the EHCS / EHS survey, using data particular to that household. Taken together, the survey households constitute a representative sample from the national population.

In order to consider the effect of uncertainty, we need to recalculate the level of fuel poverty incorporating an amount of variability. We do this using a 'Monte Carlo' model.

This approach to the measurement of uncertainty makes use of the existing fuel poverty calculation procedure; but for each household the input data are modified by the random addition and subtraction of small amounts which are representative of the level of uncertainty in the data. Each run of the Monte Carlo model produces an estimate of the number of households in fuel poverty, which differs slightly from the previous estimate and the official level of fuel poverty. If the model is run very many times, and on each occasion different adjustments to the input data are made, a distribution of the estimates of fuel poverty is generated. As long as the adjustments to the individual input data mimic correctly the unknown uncertainty associated with each value, this distribution represents the overall uncertainty in the fuel poverty estimate or the possible range of variation about a central 'best estimate'.

The starting point for the procedure is to establish frequency distributions, for each component of the input data. These distributions represent the range of possible values that the input variables could take and the likelihood of each of these values. For many of the inputs there is little or no data on which to base these distributions, and it is inevitable that arbitrary distributions must be chosen based on expert knowledge of the raw input data and data collection procedure, and the workings of the calculation itself. This is an important consideration when interpreting the results

[1] This includes all fuel used for space heating, water heating, lighting, appliances and cooking.

of this analysis, as discussed further below. A cautious approach has been taken in this analysis, with conservative assumptions in the context of fuel poverty made throughout.

Multiple runs of the model are needed to ensure convergence[2] on a distribution of possible values. Typically this will require many thousands of model runs.

When all runs have been completed this distribution can be interrogated to determine the total combined effect the introduction of uncertainty has had upon the national estimates of fuel poverty.

The analysis has been limited to two aspects of the input data which are specific to fuel poverty: variations in the fuel price inputs, and variations in the income inputs. Variations as a result of surveyor variability are not considered as these are not specific to the modelling of fuel poverty. Results are shown below for each of the different components individually, alongside combined results for the uncertainty from both income and fuel price acting together. Results have also been split by a number of household and dwelling characteristics to show the effect of uncertainty on different groups.

.

5. Producing error distributions for each component

5.1 Uncertainty in income inputs

The EHS (and EHCS before it) collects information on incomes directly from the householder as part of the interview component of the surveys.

The questions form into three main groups:

Questions on income from employment
Questions on income from state benefits
Questions on income from savings

The data collected on each of these aspects of income has an accompanying level of uncertainty. This may result from an incorrect reporting of income by the respondent, or from any structural imprecision of the interview when collecting the required data (such imprecision usually exists to maximise response rates to the question). For example, a respondent may not be fully aware of the income of other people in the household and report incorrect information, or the data may collected in banded amounts (for example, earned income and savings – recorded in this way to maximise response rates) which will result in an inevitable loss of precision.

Errors in income are typically made up of a random error component and a reporting bias. Random errors are assumed to be distributed around the reported value,

[2] Convergence is said to occur when the mean and standard deviation of the final outputs differ between two sets of model runs by less than 1%.

whereas the reporting bias tends to act towards a lower or higher value than that reported. This analysis considers the effect of random uncertainty; specific reporting biases (for example under reporting of benefit income) are not included as they are inherent in most surveys of income, and not specific to fuel poverty modelling.

The most challenging aspect of any uncertainty analysis of this type is to populate the frequency distributions of possible values which describe the uncertainty in each element of the calculation. Information on the absolute uncertainties in reported values of income from the EHS does not exist. Some limited information on the uncertainty in incomes as collected by other large national surveys of income, however, has been published. In the absence of specific information on the EHS income questions, we use these studies as a first approximation of EHS income uncertainties.

The principal source of data we have used is an examination of the uncertainty in the Family Expenditure Survey (FES, now the Living Costs and Food Survey)[3]. This analysis compared aggregate levels of income reported by the FES to data from Administrative Records (AR) over the period 1985 to 1992. The distribution of the differences between the aggregate totals recorded by the FES and the AR in this period is of particular interest. We have used these results, as a proxy for variation between reported incomes and actual income for individual cases in the EHS, to constrain the error distributions we require for the fuel poverty uncertainty analysis. As an example of this, consider Table 1 below.

Table 1: Aggregate social security benefit receipt from FES as percentage of the National Accounts. Reproduced from Table 2.3 in "How reliable is the Family Expenditure Survey?"

Year	FES / NA (% - FES total as percentage of National Accounts total)
1985	98.1
1986	95.2
1987	95.5
1988	93.4
1989	94.0
1990	93.3
1991	93.1
1992	96.4

Between 1985 and 1992 the difference between the aggregate income from social security benefits reported by the FES and AR varied between 93.1% and 98.1%. The mean difference is 94.9%, with a standard deviation of 1.76%. This standard deviation represents approximately 1.85% of the value of the mean (94.9%). This quantity (1.85%) is known as the coefficient of variation. The larger the coefficient of variation, the greater the input uncertainty.

To allow us to consider the uncertainty in total income, we use the coefficient of variation for each element of income to construct error distributions for each

[3] "How reliable is the Family Expenditure Survey? Trends in Incomes and Expenditures over Time" eds. Banks & Preston. Institute for Fiscal Studies, 1998.

household. In this process, the coefficient of variation is first converted into a cash amount by multiplying it by the (pre-uncertainty) cash value of the component of income. This gives a standard deviation which defines the error distributions around the original calculated value. Finally, we sample from these distributions randomly as part of the Monte Carlo method.

In the uncertainty analysis of fuel poverty, we have applied the most appropriate coefficients of variation derived in this way, using data from the FES, to each of the different aspects of income reported. It is recognised that the level of variation at an individual case level within any particular survey year may be very different to the level of variation shown between aggregate totals for multiple years. Aggregate totals showing little variation may in fact mask considerable variation in the underlying components of the aggregate data. However, in the absence of anything more substantive, it seems appropriate to assume that the variation between aggregate annual totals and the underlying incomes recorded at an individual case level are equivalent.

A variety of coefficients of variation have used in this analysis. These are shown in Table 2 below.

Table 2: Coefficients of variation as used in the analysis.

Aspect of income	Coefficient of variation (%)
Income from savings	15.9
Income from employment	1.6
Income from housing benefit	8.7
Income from all other benefits	1.9
Income from other sources (including occupational pensions)	6.8

5.2 Uncertainty in fuel price inputs

The fuel prices used within the fuel poverty calculations come from two main sources:

- DECC quarterly energy prices. (for metered fuels)
- Retail prices index data (for most non-metered fuels)

Gas and electricity prices are applied within the fuel poverty calculation using a mean fuel price for each government office region and method of payment. This is a simplification of the real situation where actual fuel prices vary by both supplier and tariff within each particular region. This introduces a further amount of uncertainty.

To investigate the effect of this uncertainty we are able to use supplementary data from DECC on the spread of fuel prices paid by households across the country. This additional information is based on the DECC Domestic Fuels Inquiry and provides details on the shape of the fuel price distribution for all regions and methods of payment for gas and electricity. This can be used to approximate a simple error distribution.

Table 3 below gives an example of the additional data provided, which are used to constrain the frequency distributions – this is expanded upon in the annex to this report.

Table 3: Example of distribution of average bills (for three regions and one method of payment) used to constrain the fuel price error distributions. Data of this type was used for three methods of payment and across all electricity supply regions and gas distribution zones.

Distribution zone	Method of payment	5th percentile	25th percentile	Median	75th percentile	95th percentile
N Thames	Credit	477.59	520.00	541.08	553.10	676.62
North West	Credit	477.59	526.33	553.10	553.10	635.66
North	Credit	477.59	520.99	553.10	553.10	643.69

For non-metered fuels we are able to make use of similar data on the spread of these fuels, as published as part of the Retail Price Index, in order to produce error distributions.

5.3 The adjusted fuel poverty ratio

Finally, by combining the adjusted fuel costs with the adjusted full income, a new fuel poverty ratio (FPR) for the household is calculated (see Equation 1). By counting the number of households with an FPR below 0.1 the new national estimates of fuel poverty are calculated for each model run.

6. Results

Results have been produced for uncertainty in incomes and fuel prices independently, and for the combined effect of both. As well as the main results which outline the effect upon the headline fuel poverty indicator, additional splits indicating the effect upon different types of household and dwelling have also been produced.

6.1 Uncertainty in income

The first results describe the effect of uncertainty in income when applied individually. The uncertainty analysis produces a distribution of possible values, which we can inspect to provide details on the overall uncertainty around the original estimate of fuel poverty. The distribution of possible values is shown in chart 1 below. The distribution is normally distributed and is not skewed. The ‘most likely value’ (in the middle of the distribution) is ~3.338 million fuel poor households, with a 95% confidence range between ~3.308 million and ~3.369 million households (a range of approximately 61,000 households).

Chart 1: Distribution of possible values following uncertainty in income only.

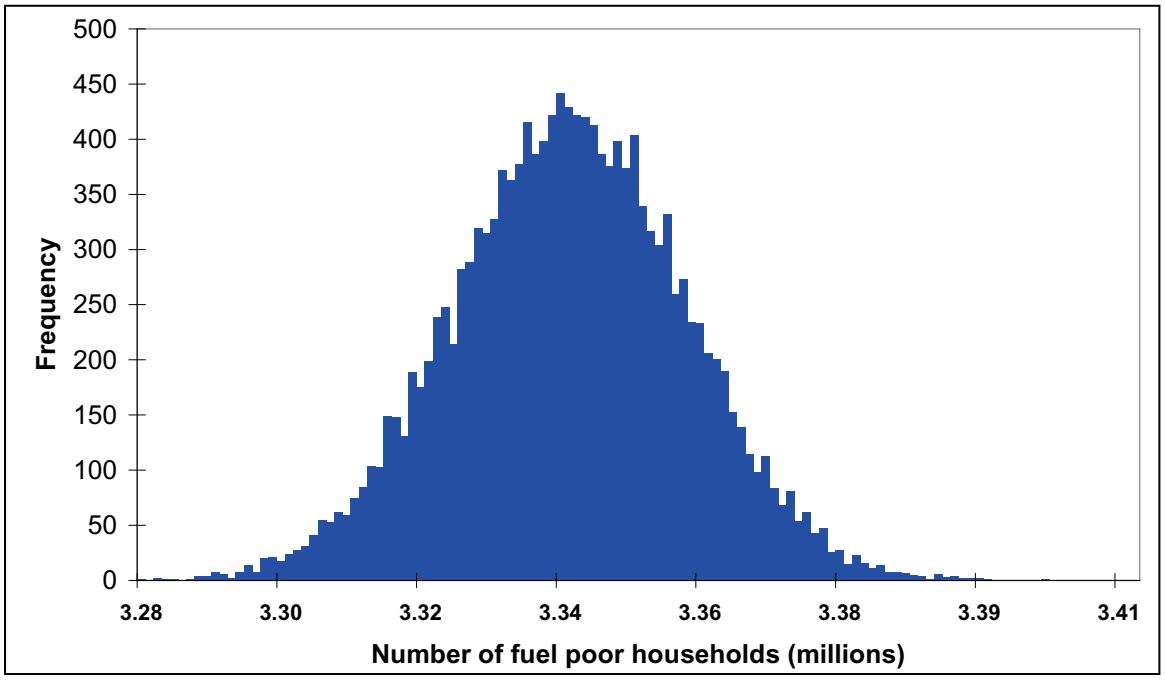


Table 4 describes the effect of uncertainty on specific household and dwelling characteristics. It shows the original values before uncertainty, the most likely values after uncertainty, the 95% confidence interval around the most likely value and the relative size of interval as a percentage of all households.

Table 4: Effect of uncertainty in incomes, split by dwelling & household characteristics (thousands of households).

	Total number of households (000s)	*Estimated number of fuel poor before uncertainty*	*Most likely value after addition of uncertainty*	*Bottom of 95% confidence interval*	*Top of 95% confidence interval*	*Width of 95% confidence interval (000s)*	*Width of interval as percentage of total number households*
All households	21,407	3,335	3,338	3,308	3,369	61	0.3%
Vulnerable households	15,130	2,650	2,662	2,634	2,689	55	0.4%
Owner occupied	14,628	2,107	2,113	2,088	2,138	50	0.3%
Private rented	2,996	600	593	580	607	26	0.9%
Local authority	1,908	375	375	366	384	18	0.9%
RSL	1,875	253	257	249	265	16	0.9%
Couple no dependent child(ren) under 60	4,085	225	223	217	230	13	0.3%
Couple no dependent child(ren) aged 60 or over	3,414	502	510	498	523	25	0.7%
Couple with dependent child(ren)	4,641	329	329	321	337	16	0.3%
Lone parent with dependent child(ren)	1,405	233	233	225	241	16	1.1%
Other multi-person household	1,640	233	228	218	238	19	1.2%
One person under 60	2,846	685	684	671	697	26	0.9%
One person 60 or over	3,376	1,127	1,132	1,113	1,151	38	1.1%
Working	12,843	851	847	832	861	28	0.2%
Unemployed	558	236	237	230	244	14	2.5%
Other inactive	8,006	2,248	2,255	2,229	2,281	53	0.7%
Household on means tested bens or tax credits with income under £15460	5,624	1,946	1,944	1,923	1,964	41	0.7%
Household not on means tested bens or tax credits with income under £15460	15,783	1,389	1,395	1,372	1,417	45	0.3%
Lowest 30% of income	6,502	2,971	2,976	2,947	3,004	57	0.9%
Highest 70% of income	14,906	364	363	351	374	24	0.2%
SAP < 30	1,550	727	723	712	734	22	1.4%
SAP 30-50	7,433	1,402	1,409	1,389	1,428	39	0.5%
SAP 51-70	10,716	1,138	1,139	1,119	1,159	39	0.4%
SAP > 70	1,708	69	67	61	74	13	0.8%
Fuel cost < £500 per yr	158	3	3	3	4	0	0.0%
Fuel cost £500-£1000 per yr	7,830	723	727	709	745	36	0.5%
Fuel cost £1000-£1500 per yr	9,353	1,496	1,501	1,479	1,523	44	0.5%
Fuel cost £1500-£2000 per yr	2,867	681	673	663	683	20	0.7%
Fuel cost £2000-£2500 per yr	730	218	221	216	226	10	1.4%
Fuel cost > £2500 per yr	469	213	213	208	218	10	2.1%

6.2 Uncertainty in fuel prices

The distribution of possible values produced for fuel prices, when this aspect of uncertainty is considered individually, is shown in chart 2 below. This histogram shows the distribution of possible values taking into account the income uncertainties. This distribution is also normally distributed and is not skewed. The most likely value is 3.336 million fuel poor households, with a 95% confidence range between ~3.295 million and ~3.377 million households (a range of approximately 81,000 households).

Chart 2: Distribution of possible values following uncertainty in fuel prices only.

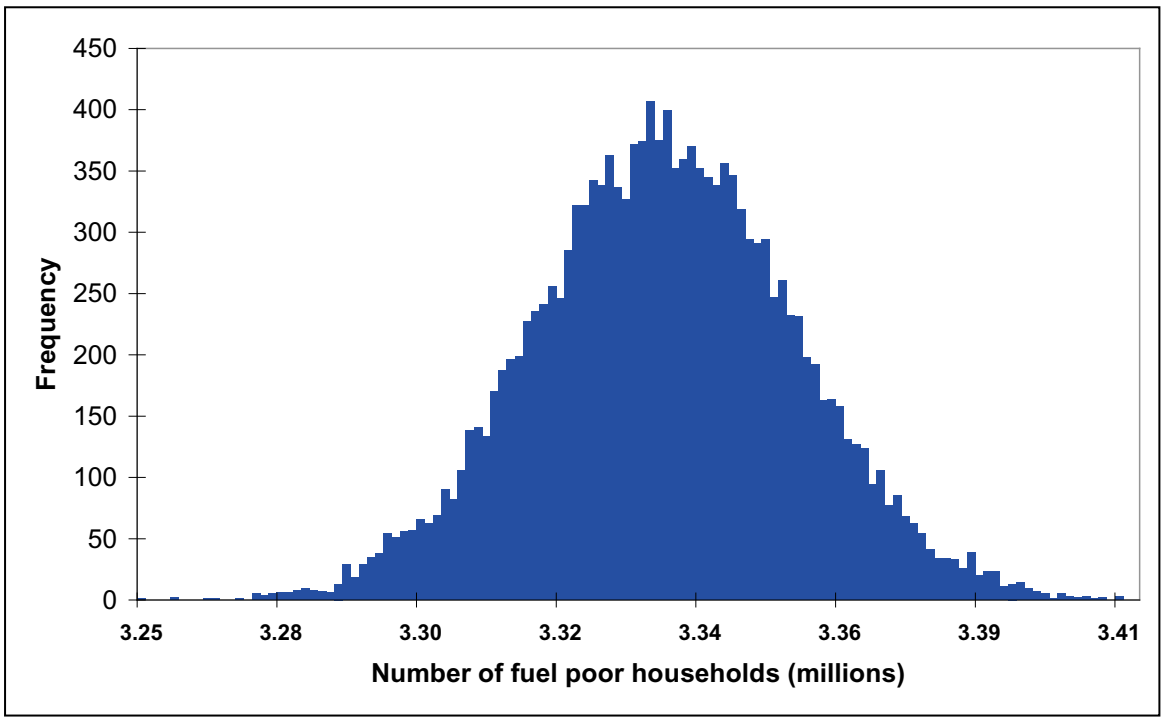


Table 5 describe the effect of uncertainty in fuel prices on specific household and dwelling characteristics. The original values before uncertainty, the most likely values after uncertainty, the 95% confidence interval around the most likely and the relative size of interval as a percentage of all households are shown in the table.

Table 5: Effect of uncertainty in fuel prices, split by dwelling & household characteristics (thousands of households).

	Total number of households	*Estimated number of fuel poor before uncertainty*	*Most likely value after addition of uncertainty*	*Bottom of 95% confidence interval*	*Top of 95% confidence interval*	*Width of 95% confidence interval*	*Width of interval as percentage of total number households*
All households	21,407	3,335	3,336	3,295	3,377	81	0.4%
Vulnerable households	15,130	2,650	2,657	2,621	2,694	74	0.5%
Owner occupied	14,628	2,107	2,121	2,086	2,156	70	0.5%
Private rented	2,996	600	594	578	609	31	1.0%
Local authority	1,908	375	365	355	374	19	1.0%
RSL	1,875	253	256	248	265	17	0.9%
Couple no dependent child(ren) under 60	4,085	225	224	213	234	21	0.5%
Couple no dependent child(ren) aged 60 or over	3,414	502	517	499	534	35	1.0%
Couple with dependent child(ren)	4,641	329	331	318	344	26	0.6%
Lone parent with dependent child(ren)	1,405	233	231	221	242	21	1.5%
Other multi-person household	1,640	233	227	216	238	22	1.3%
One person under 60	2,846	685	684	669	698	29	1.0%
One person 60 or over	3,376	1,127	1,123	1,098	1,147	49	1.5%
Working	12,843	851	855	832	878	46	0.4%
Unemployed	558	236	234	227	241	14	2.5%
Other inactive	8,006	2,248	2,247	2,214	2,279	65	0.8%
Household on means tested bens or tax credits with income under £15460	5,624	1,946	1,937	1,911	1,964	53	0.9%
Household not on means tested bens or tax credits with income under £15460	15,783	1,389	1,398	1,368	1,429	61	0.4%
Lowest 30% of income	6,502	2,971	2,963	2,928	2,999	71	1.1%
Highest 70% of income	14,906	364	372	353	392	40	0.3%
SAP < 30	1,550	727	720	703	738	35	2.3%
SAP 30-50	7,433	1,402	1,410	1,384	1,436	52	0.7%
SAP 51-70	10,716	1,138	1,139	1,114	1,164	50	0.5%
SAP > 70	1,708	69	66	60	73	13	0.8%
Fuel cost < £500 per yr	158	3	3	3	4	1	0.6%
Fuel cost £500-£1000 per yr	7,830	723	724	702	745	43	0.5%
Fuel cost £1000-£1500 per yr	9,353	1,496	1,502	1,474	1,530	56	0.6%
Fuel cost £1500-£2000 per yr	2,867	681	669	654	684	30	1.0%
Fuel cost £2000-£2500 per yr	730	218	3,336	3,295	3,377	81	11.1%
Fuel cost > £2500 per yr	469	213	2,657	2,621	2,694	74	15.8%

6.3 Combined effect of uncertainty in income and fuel prices

The distribution of possible values which is obtained when the uncertainty in income and fuel prices are combined together is shown in chart 3 below. The distribution is also normally distributed and is not skewed. The mean value for the level of fuel poverty is approximately 3.343 million households following the addition of uncertainty, with a 95% confidence interval of between 3.299 and 3.388 million households (a range of approximately 88,000 households).

As with the previous analyses of incomes and fuel prices applied in isolation, we have also produced estimates of the effect of uncertainty in these data inputs when applied together on a variety of demographic and dwelling characteristics. These are shown in Table 6.

Chart 3: Distribution of possible values following uncertainty in fuel prices and income.

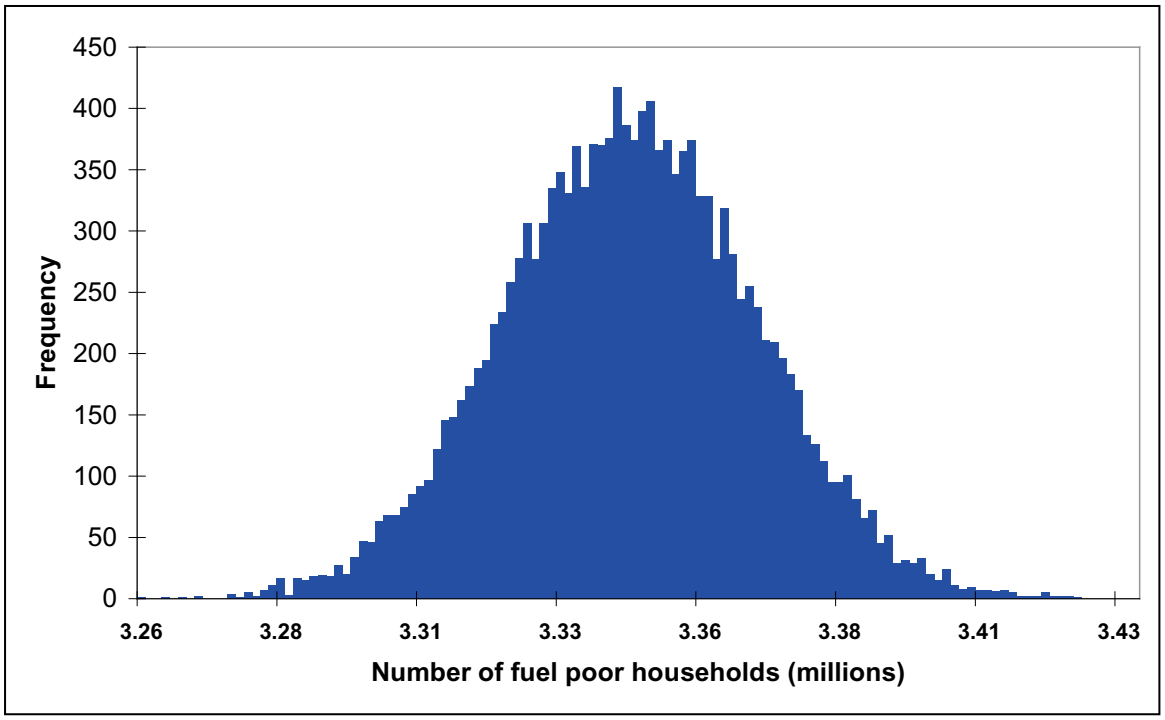


Table 6: Effect of uncertainty in both fuel prices and incomes, split by dwelling & household characteristics (thousands of households).

	Total number of households	*Estimated number of fuel poor before uncertainty*	*Most likely value after addition of uncertainty*	*Bottom of 95% confidence interval*	*Top of 95% confidence interval*	*Width of 95% confidence interval*	*Width of interval as percentage of total number households*
All households	21,407	3,335	3,343	3,299	3,388	88	0.4%
Vulnerable households	15,130	2,650	2,665	2,625	2,705	80	0.5%
Owner occupied	14,628	2,107	2,127	2,089	2,165	76	0.5%
Private rented	2,996	600	593	575	610	35	1.2%
Local authority	1,908	375	366	355	377	23	1.2%
RSL	1,875	253	257	248	267	20	1.1%
Couple no dependent child(ren) under 60	4,085	225	224	212	235	22	0.5%
Couple no dependent child(ren) aged 60 or over	3,414	502	520	501	539	38	1.1%
Couple with dependent child(ren)	4,641	329	331	317	344	27	0.6%
Lone parent with dependent child(ren)	1,405	233	232	221	244	23	1.6%
Other multi-person household	1,640	233	227	214	239	25	1.5%
One person under 60	2,846	685	684	668	701	34	1.2%
One person 60 or over	3,376	1,127	1,126	1,099	1,153	54	1.6%
Working	12,843	851	856	832	880	48	0.4%
Unemployed	558	236	233	225	241	17	3.0%
Other inactive	8,006	2,248	2,254	2,217	2,290	73	0.9%
Household on means tested bens or tax credits with income under £15460	5,624	1,946	1,937	1,908	1,966	58	1.0%
Household not on means tested bens or tax credits with income under £15460	15,783	1,389	1,407	1,373	1,440	67	0.4%
Lowest 30% of income	6,502	2,971	2,968	2,929	3,007	77	1.2%
Highest 70% of income	14,906	364	375	354	397	43	0.3%
SAP < 30	1,550	727	720	702	739	37	2.4%
SAP 30-50	7,433	1,402	1,412	1,384	1,441	57	0.8%
SAP 51-70	10,716	1,138	1,144	1,116	1,171	55	0.5%
SAP > 70	1,708	69	67	59	74	15	0.9%
Fuel cost < £500 per yr	158	3	3	3	4	1	0.6%
Fuel cost £500-£1000 per yr	7,830	723	729	705	753	48	0.6%
Fuel cost £1000-£1500 per yr	9,353	1,496	1,505	1,474	1,535	62	0.7%
Fuel cost £1500-£2000 per yr	2,867	681	669	652	685	33	1.2%
Fuel cost £2000-£2500 per yr	730	218	224	214	234	20	2.7%
Fuel cost > £2500 per yr	469	213	214	204	224	20	4.3%

6.4 Discussion of results

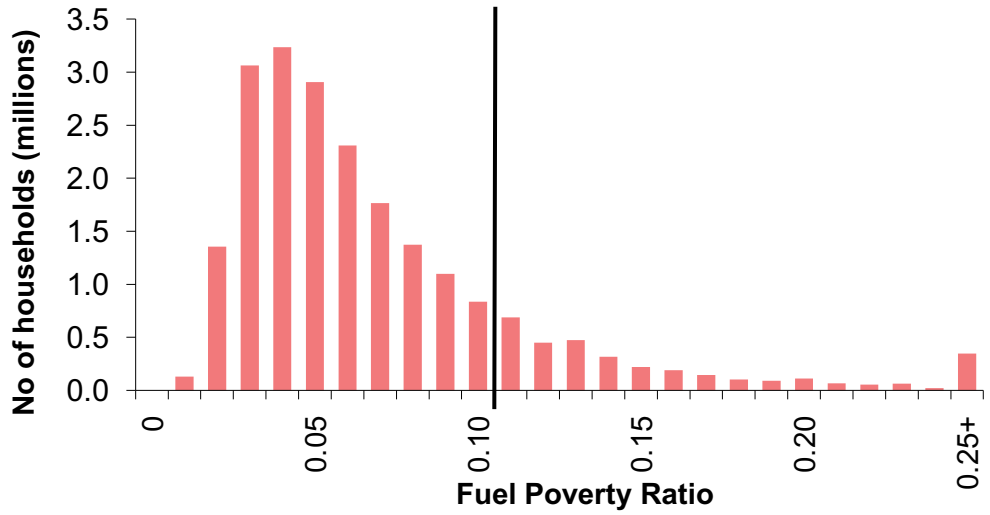
The published estimate for the number of fuel poor households in England in 2008 is approximately 3.335 million households, 2.650 million of which are vulnerable. The analysis presented above suggests that the addition of uncertainty in incomes and fuel prices is likely to increase the number of fuel poor households very slightly to a most likely value of approximately 3.343 million households, with a 95% confidence interval between 3.299 million and 3.388 million households (a range of ~88,000 households).

We can interpret this slight rise in the context of the proportion of fuel poor households on each side of the fuel poverty 10% line – i.e. how many households can be placed into fuel poverty by the addition of uncertainty, compared to how many households can be removed from fuel poverty. In particular we are interested in those close to the fuel poverty 10% threshold (as it is only these cases which could be shifted in or out of fuel poverty by the addition of uncertainty). As chart 4 shows, there are more of these households to the left of the fuel poverty line (i.e. not fuel poor) than to the right of the line (the fuel poor). This effect is clearer when considering an extreme example as given in Example 1 below.

Example 1: Uncertainty of a rare indicator

Consider a fuel poverty scenario with 100 households. Only 1 household is fuel poor and the remaining 99 are just non fuel poor. If uncertainty is introduced to the fuel poverty calculations so that it is equally likely for the FP ratio to rise or fall for each of these households, it is more likely that the level of fuel poverty will rise, as a number of the 99 non fuel poor are able to cross the line and become fuel poor whereas only 1 case could be removed from fuel poverty.

Chart 4: Distribution of households by fuel poverty ratio (before uncertainty)



The addition of uncertainty is therefore more likely to move more households into fuel poverty than it is to move out of fuel poverty. This has the net effect of increasing the average number of fuel poor households after the addition of uncertainty,

resulting in a distribution of possible values after the addition of uncertainty shifted slightly to the right of the original value. Graphically, this effect can be seen in chart 3, where the actual level of fuel poverty in 2008 (3.335m) is to the left of the mean of the distribution of fuel poverty with uncertainty simulated.

This effect was also apparent in the work undertaken on the 2006 data (see chapter 5 of http://www.decc.gov.uk/Media/viewfile.ashx?FilePath=Statistics\fuelpoverty\1_20091020153241_e_@@_annualreportfuelpovertystats2009.pdf&filetype=4&minwidth=true), although the distance between the fuel poverty level and the mean of the distribution when uncertainty is simulated was larger for the 2006 results. The "narrowing" of this gap in the 2008 results is due to a couple of reasons:

- There are over a third more fuel poor households in 2008 than in 2006
- The number of households that are marginally fuel poor (ie. with a fuel poverty ratio above 10% but below 12%) has also increased by around a quarter – these are the households that are most likely to be shifted out of fuel poverty when uncertainty is simulated on the input data..

The combination of the above means that in 2008, more households can be moved out of fuel poverty when the impacts of uncertainty are simulated than in 2006.

In this piece of work, analysis has been extended slightly towards what is known as a 'sensitivity analysis', which analyses the effect of each component of the fuel poverty calculation individually on uncertainty in number of fuel poor households. The examination of the effect of uncertainty in fuel prices and income separately suggests that the uncertainty in fuel prices has a slightly greater effect upon the level of fuel poverty than uncertainty in incomes. The 95% confidence interval around the most likely value following fuel price uncertainty is approximately 81,000 households, compared to 61,000 households for income. This suggests that efforts to reduce uncertainty may be marginally better targeted at more closely defining the fuel price inputs, rather than the income inputs.

In addition to the uncertainty resulting from the inputs directly, it is important (when considering all estimates of uncertainty) not to overlook the requirement to include the 'standard error' of the fuel poverty statistic which will further widen the confidence interval around the mean value. This results from the use of a sample of the population rather than a census. The EHS makes use of a sample of households, drawn by a random process from the whole population of households in the country. If the sample is truly representative of the total population then conclusions drawn from the sample can be attributed to the full population. There is always the chance that any sample is not perfectly representative, and this is more likely when the sample is small or has a complicated structure. The uncertainty introduced into the national estimate by the possibility of the sample not being representative is described by the standard error. The standard error is related to the sample size and the number of fuel poor households relative to the population size. Final estimates of uncertainty around the statistic should include this quantity (in addition to the uncertainty derived from the inputs directly).

These results need to be interpreted and used with caution. Any analysis of this kind is ultimately dependent upon the input distributions used within the modelling, and

the majority of the input distributions used are themselves best estimates of uncertainty for each factor. Further to this, the analysis has been designed with a cautious approach, with conservative assumptions made throughout. Uncertainty in fuel poverty estimates resulting from other factors (primarily surveyor variability) is not taken into account, nor is any bias in uncertainty estimates. As a result, these figures should be treated as indicative of the effect of input uncertainty upon national estimates of fuel poverty, rather than strictly quantitative.

Better estimates of uncertainty could be achieved by using more accurate input distributions for the input fuel poverty variables, but we are ultimately limited by the data available. To constrain the confidence range (and reduce the overall uncertainty) it may be possible to reduce the uncertainty in specific inputs. However, this is a difficult task to achieve. The EHS is already monitored and designed so as to reduce the effect of variability, and work towards this aim is ongoing. Efforts are made throughout the evolution of the survey to improve the quality of the data returned, with particular emphasis given to improving the data at source (i.e. when it is collected).

Specific improvements have been made recently in the areas of income data, and the physical data capture through a digital pen system. This has included the addition of extra cross-checks around particular aspects of income (e.g. income support), revision of the methodology for imputing incomes to align the EHS more closely to FRS income distributions and more detailed information on additional benefit units. The digital pen system was introduced in 2008 and requires online validation of survey data by the surveyors themselves, ahead of the submission of the EHS physical data. A variety of range and cross checks are built into the system for surveyors to check, which have resulted in enhanced data quality throughout the physical survey dataset.

Despite these improvements, it is essential to recognise that it is impossible to remove uncertainty completely from a survey of this type, and that the published estimates are as close to the best *measured* estimates of fuel poverty that can reasonably be achieved.

7. Summary and conclusions

The measured number of households in fuel poverty in 2008 is ~3.335 million. However, the introduction of uncertainty in incomes and fuel prices raises this value to a most likely value of around 3.343 million households, with a 95% confidence range of between ~3.299 million and ~3.388 million households.

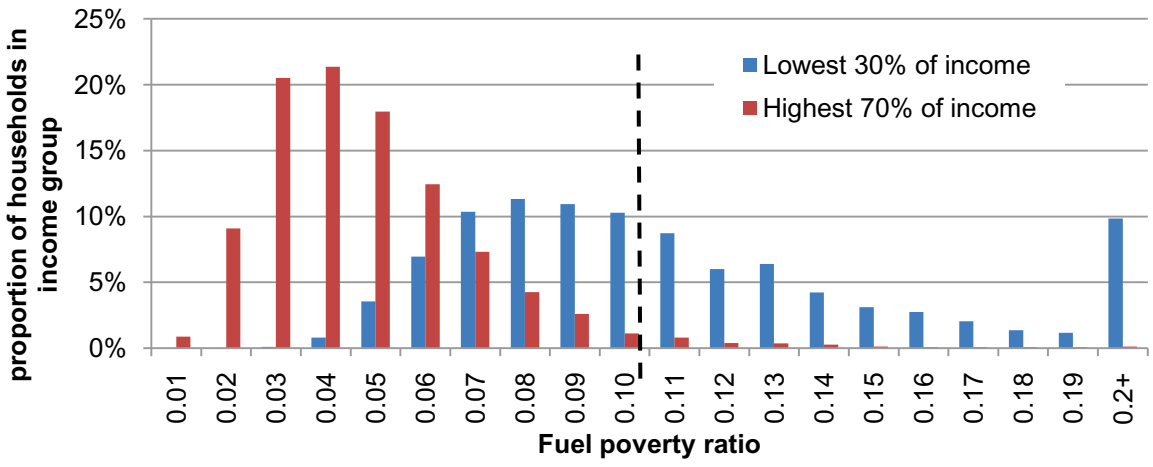
The effect of uncertainty in incomes when applied in isolation results in a confidence range around the most likely value of approximately 61,000 households, compared with approximately 81,000 households for fuel prices in isolation and approximately 88,000 when both are applied together.

8. Interpretation

As explained in section 6 of the report, the average level of fuel poverty when the input assumptions are allowed to vary to reflect potential uncertainty is greater than the official point estimate level of fuel poverty. However, the point estimate of fuel poverty **remains** the best and most accurate measure of fuel poverty.

Tables 4, 5 and 6 show how uncertainty affects different groups of households. The final column of these tables (showing the width of the 95% confidence interval as a percentage of the number of households in each group) is particularly interesting as it gives an idea of how sensitive households in each group are to the effects of uncertainty. The higher the percentage, the wider the confidence interval (relative to the size of the group), and the more likely the introduction of uncertainty is to move households into or out of fuel poverty. There are some underlying trends in these tables – for example, household groups that tend to be grouped further from the fuel poverty threshold will have smaller confidence intervals – ie. they are less likely to be affected by the impact of uncertainty compared with groups that are more spread, or located closer to the fuel poverty threshold. To illustrate this, consider household income – in the top 7 income decile groups, the width of the confidence interval is only 24,000 households, that is to say that when simulating uncertainty for these households, we can be 95% sure that the result will lie within 13,000 households below the official point estimate and around 10,000 households above it. For those households in the poorest three income decile groups, the interval is 57,000 households, with a lower bound of 24,000 below the point estimate and an upper bound of 33,000 households above the point estimate. This narrower range for higher income households is because these households are less likely to be close to the fuel poverty threshold, and so are more robust to the effects of uncertainty. This is further illustrated in the chart below, which shows the proportion of households in each "fuel poverty ratio" for each of the two income groups. Clearly, a much greater proportion of low income households are located close to the fuel poverty threshold of 10%, so it is natural that varying the circumstances of these households should have a greater impact on the fuel poverty level amongst them.

Chart 5: Distribution of fuel poverty ratios by income, England 2008



Low income households are often subjected to additional quality assurance and coherence checking in the income modelling, as they are the households most likely to be receiving benefits.

A further interesting illustration is the lack of symmetry of the confidence interval about the official point estimate of fuel poverty for different groups. For example, consider again households split into the two income groups. The table below shows the impact of uncertainty due to fuel prices on the two income groups – the two columns on the far right show the boundaries of the 95% confidence interval relative to the actual level of fuel poverty for the group in 2008. The table shows that when simulating uncertainty around prices, the effect is more likely to reduce the level of fuel poverty amongst low income households (as the lower bound is further from the actual level than the upper bound), and increase the level amongst higher income households.

Table 7 - Uncertainty in fuel prices

no. of households (000s)	No. of hh fuel poor	Lower bound for CI	Upper bound for CI	Lower bound relative to actual	Upper bound relative to actual
Lowest 30% of income	2,971	2,928	2,999	-1.4%	0.9%
Highest 70% of income	364	353	392	-3.0%	7.7%

A similar comparison holds for the uncertainty in incomes.

Table 8 - Uncertainty in income

no. of households (000s)	No. of hh fuel poor	Lower bound for CI	Upper bound for CI	Lower bound relative to actual	Upper bound relative to actual
Lowest 30% of income	2,971	2,929	3,007	-1.4%	1.2%
Highest 70% of income	364	354	397	-2.7%	9.1%

However, when comparing the confidence intervals for the uncertainty in factors combined, the interactions of prices and income result in a more symmetric confidence interval.

9. Annex – input distributions

Section 4.2 of the report explains how average fuel prices are constructed from a near census of gas and electricity prices within DECC. This section summarises the price distributions, and shows how tightly prices are distributed around the median within each region and for each of the three main payment methods. In the charts, each region of the country is shown along the horizontal axis, labelled from 1 to 10. The vertical axis shows the spread of prices about the median in the region. The horizontal axis itself represents the median tariff, or price, for each region. The diamonds represent the inter-quartile range of prices – for example, around a quarter of customer within region 1 experience prices between -4% of the median and the median itself for direct debit (chart 6.1). The triangles represent the lowest 5% (green) and upper 5% (or 95%, in orange) extreme prices. Therefore, 90% of all customers within a region experience prices within this range. Prices outside of this range are not considered in this work – in some cases these can be quite extreme and may be accredited to obscure or outlying tariffs. The "prices" plotted represent a bill based on an average level of consumption for each tariff in the region and payment method, across 2008.

Gas price distributions by region and payment method

Chart 6.1 - Direct debit

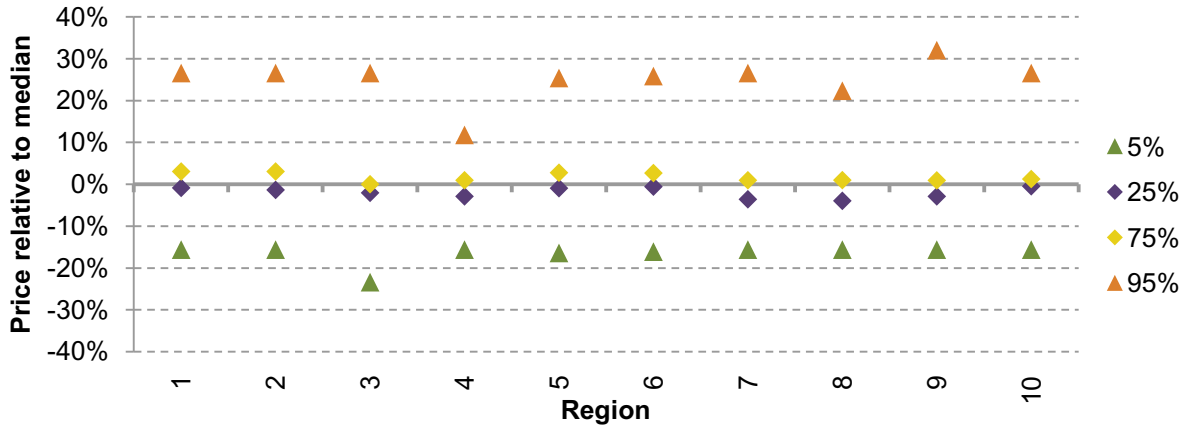


Chart 6.2 - Standard credit

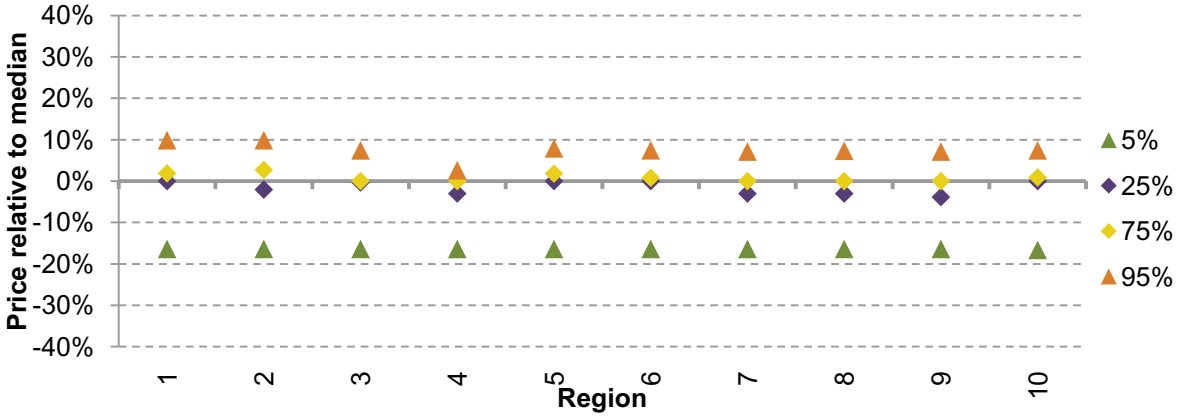
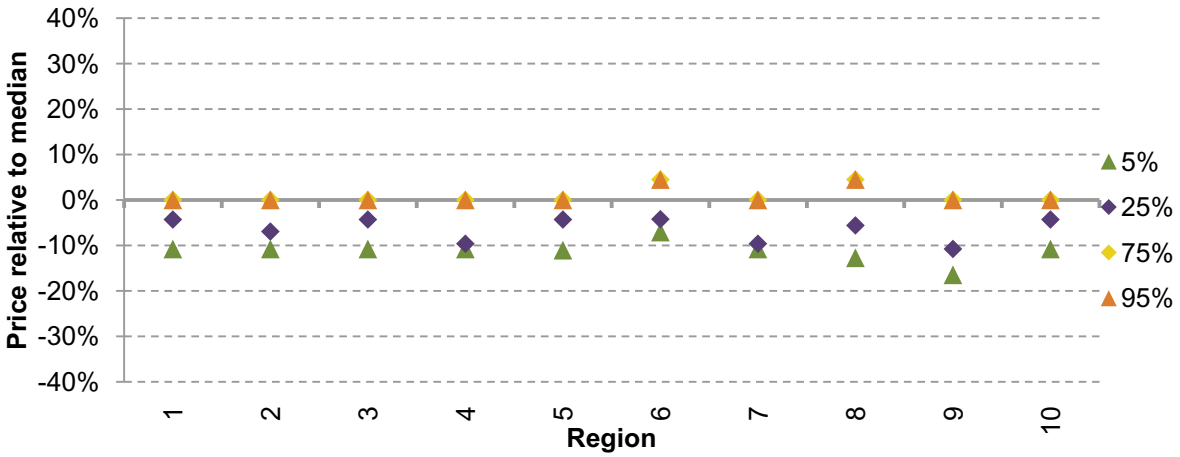


Chart 6.3 - Pre-payment meter



The charts illustrate that at least half of customers paying for their gas by either direct debit or standard credit pay within 4 per cent of the median in any region. The range is slightly wider for pre-payment meter, largely because there are a greater number of pre-payment meter tariffs just below the median level.

Electricity price distributions by region and payment method

Prices for electricity show a little more spread about the median compared with gas. However, the majority are still within 10 per cent of the median regardless of region or payment method.

Chart 7.1 - Direct debit

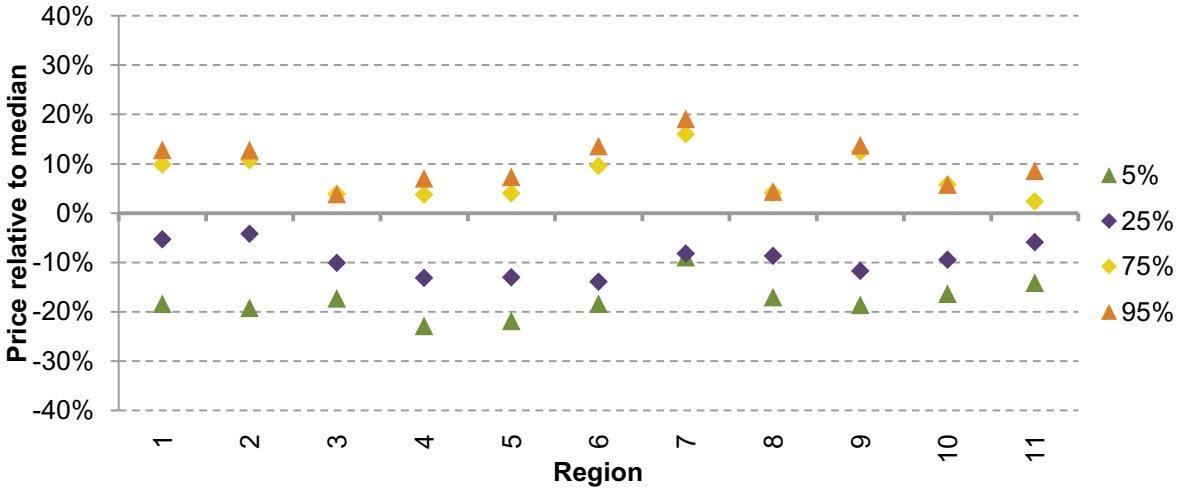


Chart 7.2 - Standard credit

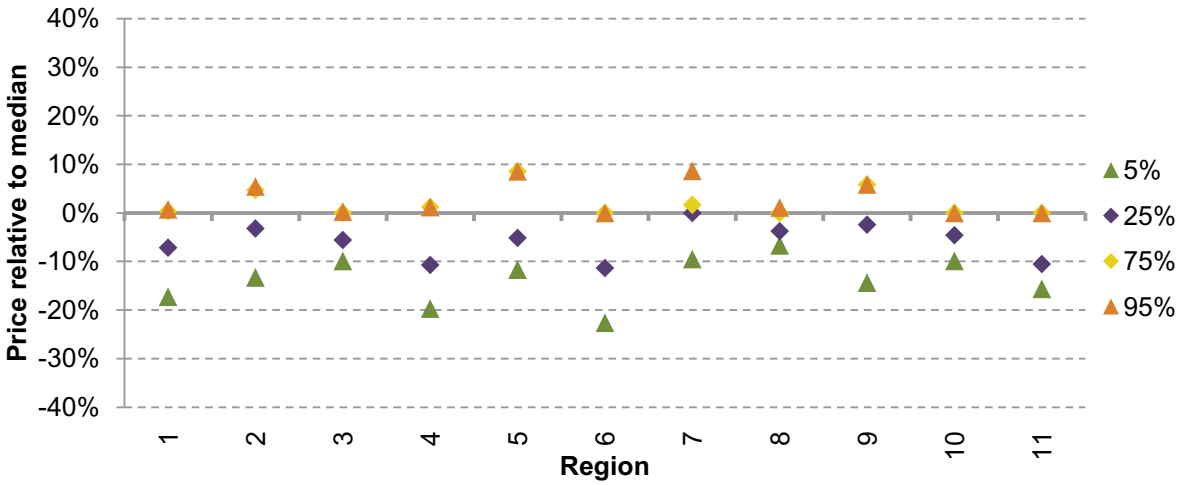
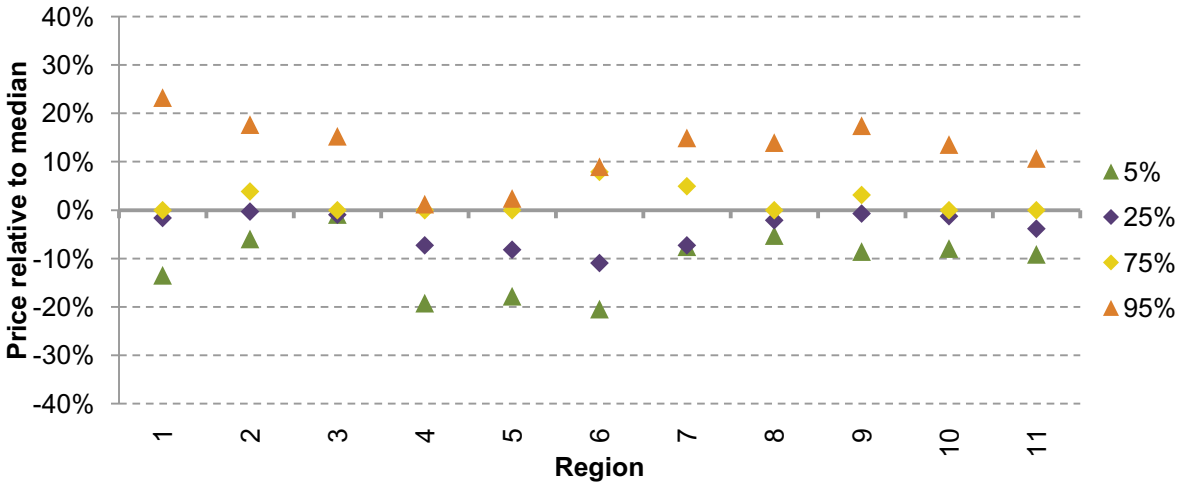


Chart 7.3 - Pre-payment meter



Uncertainty in income distributions

The chart below shows an illustrative probability distribution for benefit income, for a household reporting a benefit income of £10,000. This is estimated using the figures in table 2. It shows that there is approximately a 10% chance that benefit income is under reported by around £250 (or 2.5 per cent) – and similarly around a 10% chance that it is over reported by around £250. In reality, uncertainty around benefit income is unlikely to be normally distributed, and so this assumption is approximate in the absence of more detailed information.

Chart 8 – Probability distribution of actual benefit income, based on a reported benefit income of £10,000